Matching the proteome to the genome

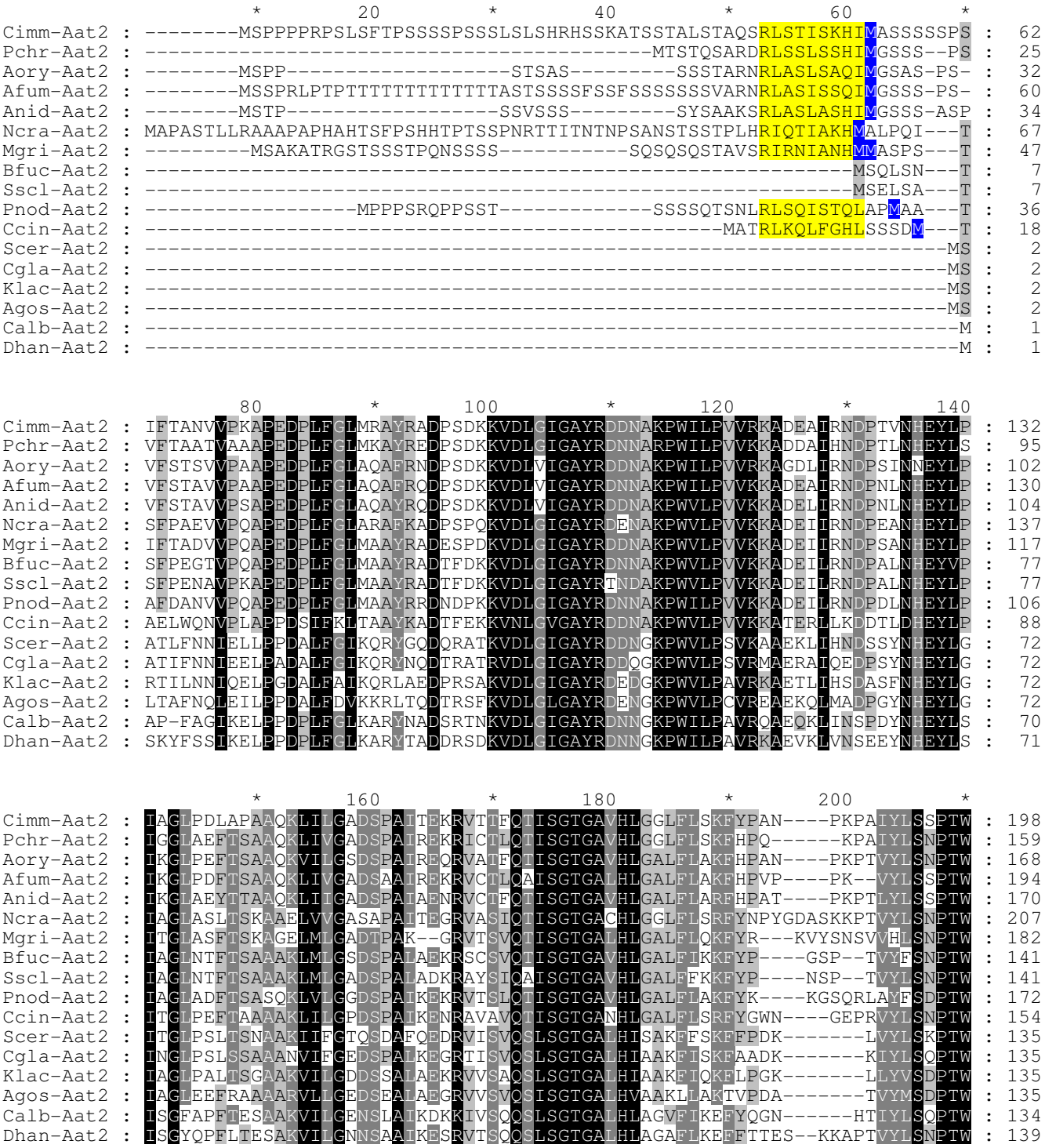
Kiel, Jan A. K. W.; van den Berg, Marco A.; Fusetti, Fabrizia; Poolman, Bert; Bovenberg, Roel A. L.; Veenhuis, Marten; van der Klei, Ida J.

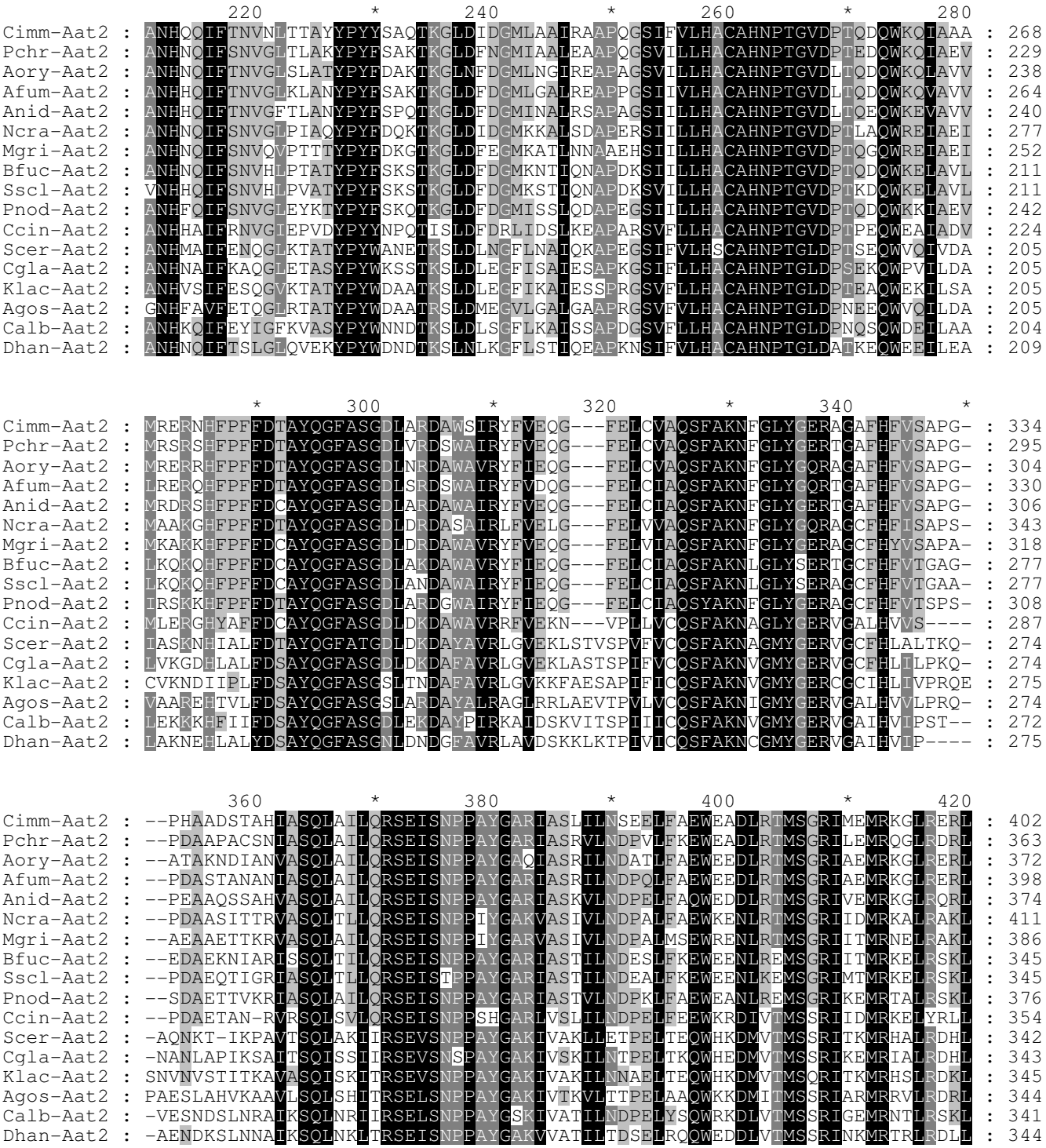
Published in:
FUNCTIONAL & INTEGRATIVE GENOMICS

DOI:
10.1007/s10142-009-0110-6

Publication date:
2009

Supplementary Fig. S1




220 * 240 * 260 * 280
Cimm-Aat2 : ANHQQIFTNVNLTTAYYPYYSAQTKGLDIDGMLAAIRAAPQGSIFVLHACAHNPTGVDPTQDQWKQIAAA : 268
Pchr-Aat2 : ANHNQIFSNVGLTLAKYPYFSAKTKGLDFNGMIAALEAAPQGSVILLHACAHNPTGVDPTEDQWKQIAEV : 229
Aory-Aat2 : ANHNQIFTNVGLSLATYPYFDAKTKGLNFDGMLNGIREAPAGSVILLHACAHNPTGVDLTQDQWKQLAVV : 238
Afum-Aat2 : ANHHQIFTNVGLKLANYPYFSAKTKGLDFDGMLGALREAPPGSIIVLHACAHNPTGVDLTQDQWKQVAVV : 264
Anid-Aat2 : ANHHQIFTNVGFTLANYPYFSPQTKGLDFDGMINALRSAPAGSIILLHACAHNPTGVDLTQEQWKEVAVV : 240
Ncra-Aat2 : ANHNQIFSNVGLPIAQYPYFDQKTKGLDIDGMKKALSDAPERSIILLHACAHNPTGVDPTLAQWREIAEI : 277
Mgri-Aat2 : ANHNQIFSNVQVPTTTYPYFDKGTKGLDFEGMKATLNNAAEHSIILLHACAHNPTGVDPTQGQWREIAEI : 252
Bfuc-Aat2 : ANHNQIFSNVHLPTATYPYFSKSTKGLDFDGMKNTIQNAPDKSIILLHACAHNPTGVDPTQDQWKELAVL : 211
Sscl-Aat2 : VNHHQIFSNVHLPVATYPYFSKSTKGLDFDGMKSTIQNAPDKSVILLHACAHNPTGVDPTKDQWKELAVL : 211
Pnod-Aat2 : ANHFQIFSNVGLEYKTYPYFSKQTKGLDFDGMISSLQDAPEGSIILLHACAHNPTGVDPTQDQWKKIAEV : 242
Ccin-Aat2 : ANHHAIFRNVGIEPVDYPYYNPQTISLDFDRLIDSLKEAPARSVFLLHACAHNPTGVDPTPEQWEAIADV : 224
Scer-Aat2 : ANHMAIFENQGLKTATYPYWANETKSLDLNGFLNAIQKAPEGSIFVLHSCAHNPTGLDPTSEQWVQIVDA : 205
Cgla-Aat2 : ANHNAIFKAQGLETASYPYWKSSTKSLDLEGFISAIESAPKGSIFLLHACAHNPTGLDPSEKQWPVILDA : 205
Klac-Aat2 : ANHVSIFESQGVKTATYPYWDAATKSLDLEGFIKAIESSPRGSVFLLHACAHNPTGLDPTEAQWEKILSA : 205
Agos-Aat2 : GNHFAVFETQGLRTATYPYWDAATRSLDMEGVLGALGAAPRGSVFVLHACAHNPTGLDPNEEQWVQILDA : 205
Calb-Aat2 : ANHKQIFEYIGFKVASYPYWNNDTKSLDLSGFLKAISSAPDGSVFLLHACAHNPTGLDPNQSQWDEILAA : 204
Dhan-Aat2 : ANHNQIFTSLGLQVEKYPYWDNDTKSLNLKGFLSTIQEAPKNSIFVLHACAHNPTGLDATKEQWEEILEA : 209
* 300 * 320 * 340 *
Cimm-Aat2 : MRERNHFPFFDTAYQGFASGDLARDAWSIRYFVEQG---FELCVAQSFAKNFGLYGERAGAFHFVSAPG- : 334
Pchr-Aat2 : MRSRSHFPFFDTAYQGFASGDLVRDSWAIRYFVEQG---FELCVAQSFAKNFGLYGERTGAFHFVSAPG- : 295
Aory-Aat2 : MRERRHFPFFDTAYQGFASGDLNRDAWAVRYFIEQG---FELCVAQSFAKNFGLYGQRAGAFHFVSAPG- : 304
Afum-Aat2 : LRERQHFPFFDTAYQGFASGDLSRDSWAIRYFVDQG---FELCIAQSFAKNFGLYGQRTGAFHFVSAPG- : 330
Anid-Aat2 : MRDRSHFPFFDCAYQGFASGDLARDAWAIRYFVEQG---FELCIAQSFAKNFGLYGERTGAFHFVSAPG- : 306
Ncra-Aat2 : MAAKGHFPFFDTAYQGFASGDLDRDASAIRLFVELG---FELVVAQSFAKNFGLYGQRAGCFHFISAPS- : 343
Mgri-Aat2 : MKAKKHFPFFDCAYQGFASGDLDRDAWAVRYFVEQG---FELVIAQSFAKNFGLYGERAGCFHYVSAPA- : 318
Bfuc-Aat2 : LKQKQHFPFFDCAYQGFASGDLAKDAWAVRYFIEQG---FELCIAQSFAKNLGLYSERTGCFHFVTGAG- : 277
Sscl-Aat2 : LKQKQHFPFFDCAYQGFASGDLANDAWAIRYFIEQG---FELCIAQSFAKNLGLYSERAGCFHFVTGAA- : 277
Pnod-Aat2 : IRSKKHFPFFDTAYQGFASGDLARDGWAIRYFIEQG---FELCIAQSYAKNFGLYGERAGCFHFVTSPS- : 308
Ccin-Aat2 : MLERGHYAFFDCAYQGFASGDLDKDAWAVRRFVEKN---VPLLVCQSFAKNAGLYGERVGALHVVS---- : 287
Scer-Aat2 : IASKNHIALFDTAYQGFATGDLDKDAYAVRLGVEKLSTVSPVFVCQSFAKNAGMYGERVGCFHLALTKQ- : 274
Cgla-Aat2 : LVKGDHLALFDSAYQGFASGDLDKDAFAVRLGVEKLASTSPIFVCQSFAKNVGMYGERVGCFHLILPKQ- : 274
Klac-Aat2 : CVKNDIIPLFDSAYQGFASGSLTNDAFAVRLGVKKFAESAPIFICQSFAKNVGMYGERCGCIHLIVPRQE : 275
Agos-Aat2 : VAARERHTVLFDSAYQGFASGSLARDAYALRAGLRRLAEVTPVLVCQSFAKNIGMYGERVGALHVVLPRQ- : 274
Calb-Aat2 : LEKKKHFIIFDSAYQGFASGDLEKDAYPIRKAIDSKVITSPIIICQSFAKNVGMYGERVGAIHVIPST-- : 272
Dhan-Aat2 : LAKNEHLALYDSAYQGFASGNLDNDGFAVRLAVDSKKLKTPIVICQSFAKNCGMYGERVGAIHVIP---- : 275
360 * 380 * 400 * 420
Cimm-Aat2 : --PHAADSTAHIASQLAILQRSEISNPPAYGARIASLILNSEELFAEWEADLRTMSGRIMEMRKGLRERL : 402
Pchr-Aat2 : --PDAAPACSNIASQLAILQRSEISNPPAYGARIASRVLNDPVLFKEWEADLRTMSGRILEMRQGLRDRL : 363
Aory-Aat2 : --ATAKNDIANVASQLAILQRSEISNPPAYGAQIASRILNDATLFAEWEEDLRTMSGRIAEMRKGLRERL : 372
Afum-Aat2 : --PDASTANANIASQLAILQRSEISNPPAYGARIASRILNDPQLFAEWEEDLRTMSGRIAEMRKGLRERL : 398
Anid-Aat2 : --PEAAQSSAHVASQLAILQRSEISNPPAYGARIASKVLNDPELFAQWEDDLRTMSGRIVEMRKGLRQRL : 374
Ncra-Aat2 : --PDAASITTRVASQLTLLQRSEISNPPIYGAKVASIVLNDPALFAEWKENLRTMSGRIIDMRKALRAKL : 411
Mgri-Aat2 : --AEAAETTKRVASQLAILQRSEISNPPIYGARVASIVLNDPALMSEWRENLRTMSGRIITMRNELRAKL : 386
Bfuc-Aat2 : --EDAEKNIARISSQLTILQRSEISNPPAYGARIASTILNDESLFKEWEENLREMSGRIITMRKELRSKL : 345
Sscl-Aat2 : --PDAEQTIGRIASQLTLLQRSEISTPPAYGARIASTILNDEALFKEWEENLKEMSGRIMTMRKELRSKL : 345
Pnod-Aat2 : --SDAETTVKRIASQLAILQRSEISNPPAYGARIASTVLNDPKLFAEWEANLREMSGRIKEMRTALRSKL : 376
Ccin-Aat2 : --PDAETAN-RVRSQLSVLQRSEISNPPSHGARLVSLILNDPELFEEWKRDIVTMSSRIIDMRKELYRLL : 354
Scer-Aat2 : -AQNKT-IKPAVTSQLAKIIRSEVSNPPAYGAKIVAKLLETPELTEQWHKDMVTMSSRITKMRHALRDHL : 342
Cgla-Aat2 : -NANLAPIKSAITSQISSIIRSEVSNSPAYGAKIVSKILNTPELTKQWHEDMVTMSSRIKEMRIALRDHL : 343
Klac-Aat2 : SNVNVSTITKAVASQISKITRSEVSNPPAYGAKIVAKILNNAELTEQWHKDMVTMSQRITKMRHSLRDKL : 345
Agos-Aat2 : PAESLAHVKAAVLSQLSHITRSELSNPPAYGAKIVTKVLTTPELAAQWKKDMITMSSRIARMRRVLRDRL : 344
Calb-Aat2 : -VESNDSLNRAIKSQLNRIIRSELSNPPAYGSKIVATILNDPELYSQWRKDLVTMSSRIGEMRNTLRSKL : 341
Dhan-Aat2 : -AENDKSLNNAIKSQLNKLTRSEISNPPAYGAKVVATILTDSELRQQWEDDLVTMSSRINKMRTRLRDLL : 344

```
                        *       440         *       460         *       480         *
Cimm-Aat2 : EQK-GTPGTWEHITNQIGMFSFTGLTEQQVGILREKWHVYMTKNGRISMAGLNTHNIDYFAEAVDSVVRE : 471
Pchr-Aat2 : EKR-GTPGSWDHITSQIGMFSFTGLSEEQVLTLRSKWHVYMTKNGRISMAGLNTNNIDYFAEAVDSVVRE : 432
Aory-Aat2 : EAK-GTPGTWNHVTDQIGMFSFTGLSEAQVKLLREKWHIYMTKNGRISMAGLNTHNIDYFAEAVDSVVRE : 441
Afum-Aat2 : EAK-GTPGNWEHITSQIGMFSFTGLTEEQVKILREKWHVYMTKNGRISMAGLNTHNLDYFAEAVDSVVRE : 467
Anid-Aat2 : EEK-KTPGSWNHITDQIGMFSFTGLSEAQVKVLREKWHVYMTKNGRISMAGLNSHNLDYFAEAVDSVVRE : 443
Ncra-Aat2 : EEL-GTPGPWNHITDQIGMFSFTGLNEKQVAKLREEFHIYMTKNGRISMAGLNKKNVDYVARAVDKVVRE : 480
Mgri-Aat2 : EAL-GTPGTWNHITDQIGMFSFTGLTEAQVLKIRSDYHVYMTKNGRISMAGLNSRNIDYFAKAVDKVVRE : 455
Bfuc-Aat2 : EEM-GTPGKWNHITEQIGMFSFTGLSEKQVMELRDEAHVYMTKNGRISMAGLNTNNIDYFARAVDKVVRK : 414
Sscl-Aat2 : EEM-GTPGKWNHITDQIGMFSFTGLTEKQVLELRSEAHVYMTNNGRISMAGLNTNNIGYFAQAVDKVVRK : 414
Pnod-Aat2 : EEL-GTPGTWNHITDQIGMFSFTGLTEQQVLKIREDAHVYMTKNGRISMAGLNTNNVEYFAKAVDKVVRE : 445
Ccin-Aat2 : TEVYKTPGNWDHILNQIGMFSFTGLNPEQSKAMVEKAHIYMTGNGRISMAGLNSHNIEYFASSLDKAVRG : 424
Scer-Aat2 : VK-LGTPGNWDHIVNQCGMFSFTGLTPQMVKRLEETHAVYLVASGRASIAGLNQGNVEYVAKAIDEVVRF : 411
Cgla-Aat2 : VK-LGTPGTWDHIVEQCGMFSFTGLTPEMVKRLEVQHAVYLVSSGRASIAGLNSGNVEYTAKAIDEVVRH : 412
Klac-Aat2 : LE-LQTPGNWDHIVNQCGMFSYTGLTKPMVERLEKEHGVYMVASGRASIAGLNEKNVEHVAKAINEVVRH : 414
Agos-Aat2 : VE-LGTPGNWDHIVQQCGMFSYTGLTKEMVARMEKDFAIYMVSSGRISIAGLNDSNVGHVANAIDNAVRH : 413
Calb-Aat2 : ES-LGTPGTWNHITEQTGMFSFTGLTPQMVERLEKHHGIYLVSSGRASVAGLNEHNVDQVAKAIDEVVRH : 410
Dhan-Aat2 : TNQYSTPGNWDHIVSQSGMFSFTGLTPDMVSRLEKNHGIYLVSSGRASVAGLNDGNVEQVAKAIDEVVRF : 414


Cimm-Aat2 : TSS----- : 474
Pchr-Aat2 : SAKL---- : 436
Aory-Aat2 : TS------ : 443
Afum-Aat2 : TS------ : 469
Anid-Aat2 : TS------ : 445
Ncra-Aat2 : VQ------ : 482
Mgri-Aat2 : TQ------ : 457
Bfuc-Aat2 : TQQESAHL : 422
Sscl-Aat2 : TQQETANL : 422
Pnod-Aat2 : TQ------ : 447
Ccin-Aat2 : EL------ : 426
Scer-Aat2 : YTIEAKL- : 418
Cgla-Aat2 : YSTSKL-- : 418
Klac-Aat2 : FNESSKL- : 421
Agos-Aat2 : FSTPKL-- : 419
Calb-Aat2 : FGKSKL-- : 416
Dhan-Aat2 : YSKSKL-- : 420
```

Legend to Supplementary Fig. S1.

Sequence alignment of fungal microbody-located aspartate transaminases (Aat2).

The following accession numbers were used:

Fungi:

Cimm, *Coccidioides immitis* (EAS29116-corrected); Pchr, *Penicillium chrysogenum* (Pc22g19440); Aory, *Aspergillus oryzae* (BAE65140); Afum, *Aspergillus fumigatus* (EAL93260); Anid, *Aspergillus nidulans* (EAA58023); Ncra, *Neurospora crassa* (EAA33221); Mgri, *Magnaporthe grisea* (EDJ99457); Bfuc, *Botryotinia fuckeliana* (EDN20556); Sscl, *Sclerotinia sclerotiorum* (EDN99237); Pnod, *Phaeosphaeria nodorum* (translation of AAGI01000295, nt 39726-38235); Ccin, *Coprinopsis cinerea* (EAU88611).

Yeast species:

Scer, *Saccharomyces cerevisiae* (NP_013127); Cgla, *Candida glabrata* (CAG58407); Klac, *Kluyveromyces lactis* (CAG98584); Agos, *Ashbya gossypii* (AAS53582); Calb, *Candida albicans* (EAK91905); Dhan, *Debaryomyces hansenii* (CAG87700).

Sequences were aligned using the Clustal_X programme (Thompson et al. 1997). Gaps were introduced to maximize the similarity. Residues that are similar in all 17 proteins are represented by white letters that are shaded black. Similar residues in at least 14 of the proteins are shown as white letters that are shaded dark grey, while those that are similar in at least 11 of the proteins are shown as black letters that are shaded light grey. Putative PTS1 sequences are shaded red; putative PTS2 sequences are in yellow, while the Met residue that immediately follows the putative PTS2 sequences is shown in white on a blue background.

Reference: (for Figs S1, S2 and S3)

1. Thompson JD, Gibson TJ, Plewniak F et al (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 24:4876-4882